République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université 8 Mai 1945 Guelma

Faculté des Mathématiques et de l'Informatique et des Sciences de la Matière
Département de Mathématiques

Mémoire

Présenté en vue de l'obtention du diplôme de

Master Académique en Mathématiques

Option : Probabilités et applications

Par :

Mme : SAIBI Messaouda et Melle : FEHADA Houda

Intitulé

Approche EDP classique de la programmation dynamique

Dirigé par : Dr. Debbar Rabah

Devant le jury

PRESIDENT	Mme. REBAI. G	Univ-Guelma
RAPPORTEUR	Mr. DEBBAR. R	Univ-Guelma
EXAMINATEUR1	Mr. BENCHAABANE. A	Univ-Guelma

Session Juin 2013

Remerciements

Ce mémoire n'aurait vu le jour sans les gens qui m'entourent et qui nous ont encouragés à le mener à son terme. Mes premiers remerciements, les plus vifs, vont à notre encadreur, *Mr. DEBBAR .RABAH* qui nous accompagné tout au long de cette investigation et qu'il trouver exprimée ici notre profonde gratitude et notre reconnaissance.

Nous voudrons remercier, en deuxième lieu les membres du jury : *Mme. REBAI.G* et *Mr. BENCHAABEN.A* Pour l'honneur qu'ils nous font en participant à l'évaluation de notre travail. Nous ne pourrons passer sous silence la présence, l'affection et l'inestimable soutient de notre parents et notre famille.

Un grand merci aux tous notre professeurs depuis le primaire, qui nous ont donné le gout des maths et des sciences en général.

Nous gardons toujours le meilleur souvenir de tous notre collègues de l'université de 08 Mai 45. Merci à mes collègues de promotion de probabilités pour l'ambiance chaleureuse à laquelle ils contribuent. Que soient, enfin, remerciées toutes les personnes qui ne trouvent pas leur nom sur cette page, mais qui ont aidé et concouru à la réalisation de ce mémoire.

RÉSUMÉ

Dans ce mémoire, nous étudions la méthode de la programmation dynamique pour résoudre un problème de contrôle stochastique. Le cadre adopté dans la deuxième chapitre (partie 1) est celui des processus de diffusions contrôlés à valeurs dans $\mathbb{R}^n$ et le problème formulé est en horizon fini ou en horizon infini.

L'idée basique de la méthode est de considérer une famille de problèmes de contrôle à différents états initiaux et d'établir des relations entre les fonctions valeurs associées. Ce principe, appelé principe de la programmation dynamique, est énoncé précisément dans la deuxième chapitre (partie 2).

L'équation de la programmation dynamique dans la troisième chapitre conduit à une équation aux dérivées partielles (EDP) nonlinéaire du second ordre, appelée Hamilton-Jacobi-Bellman (HJB).

Table des matières

Introduction

Dans ce travaille, nous étudions la méthode de la programmation dynamique pour résoudre un problème de contrôle stochastique. Le cadre adopté dans la deuxième chapitre est celui des processus de diffusion contrôlés à valeurs dans $\mathbb{R}^n$ et le probleme formulé est en horizon fini ou horizon infini.

L'idée basique de la méthode est de considérer une famille de problèmes de contrôle à différents états initiaux et d'établir des relations entre les fonctions valeurs associées. Ce principe, appelé *principe de la programmation dynamique* et initié dans les année 50 par Ballman . L'équation de la programmation dynamique conduit à une équations aux dérivées partielles (EDP) nonlinéaire du second ordre, appelée *Hamilton-Jacobi-Bellman (HJB).* Lorsque cette *EDP* peut être résolue par l'obtention explicite ou théorique d'une solution régulière, le théorème de vérification, valide l'optimalité de ce candidat solution de HJB, et permet aussi de caractériser un contrôle optimal. Cette approche classique de la programmation dynamique est appelée étape de vérification. Nous illustrons cette mèthode en section 3.3 (troisième chapitre) sur deux exemples de modèles en finance. l'inconvénient majeur de cette approche est de supposer l'existence d'une solution régulière à l'EDP d' HJB. Ce n'est pas toujours le cas et nous illustrons ce fait sur exemple simple en section 3.3.3 (troisième chapitre) inspiré de la finance.

Chapitre 1

Notion préliminaire

1.1 Equation différentielle stochastique

Le concept d'équation différentielle stochastique généralise celui d'équation différentielle ordinaire aux processus stochastiques. La formalisation théorique de ce problème à elle seule a posé problème aux mathématiciens, et il a fallu attendre les années 40 et les travaux du mathématicien japonais Itô Kiyoshi pour la définition de l'intégrale stochastique. Il s'agit d'étendre la notion d'intégrale de Lebesgue aux processus stochastique selon un mouvement brownien. On construira cette intégrale, et on donnera un sens à l'expression

$$\int_s^t f(u,w)dB_u$$

où $f(u,.)$ est un processus stochastique muni de propriétés de régularité suffisantes.
À partir de la théorie de l'intégration, on construit la théorie des équations différentielles stochastiques EDS.
Soit $(\Omega, \mathcal{F}, F = (\mathcal{F}_t)_{t\in[0,T]}, P)$ un espace de probabilité filtré satisfaisant les conditions habituelles, et soit $(B_t)_t$ un $\mathcal{F}_t$-mouvement brownien à valeurs dans $\mathbb{R}$. On se donne des fonctions $b(t,x,w)$ et $\sigma(t,x,w)$ définies sur $[0,T]\times\mathbb{R}\times\Omega$ à valeurs dans $\mathbb{R}$. On suppose que pour tout w, les fonctions $b(.,.,w)$ et $\sigma(.,.,w)$ sont boréliennes sur $[0,T]\times\mathbb{R}$ et pour tout $x\in\mathbb{R}$, les processus $b(.,x,.)$ et $\sigma(.,x,.)$, notés parfois pour simplifier $b(.,x)$ et $\sigma(.,x)$ sont progressifs. On considère alors l'EDS à valeurs dans $\mathbb{R}$:

$$dX_t = b(t,X_t)dt + \sigma(t,X_t)dB_t. \tag{1.1}$$

Les coefficients $b(t,x,w)$ et $\sigma(t,x,w)$ sont de la forme $\tilde{b}(t,x,\alpha_t(w))$, $\tilde{\sigma}(t,x,\alpha_t(w))$ où $\tilde{b},\tilde{\sigma}$

sont des fonctions déterministes boréliennes sur $[0,T] \times \mathbb{R} \times A$, A sous ensemble de $\mathbb{R}$, et $\alpha = (\alpha_t)_{t\in[0,T]}$ est un processus progressif à valeurs dans A.

Une solution de (1.1) porte le nom de "diffusion". Ces équations permetent de construire la plupart des modèles d'actifs utiles en finances, aussi bien lorsque l'on cherche à modéliser des actifs des taux d'intérêt.

Précisons, ce que l'on entend par une solution forte de l'EDS (1.1).

Définition 1.1.1. (Solution forte d'EDS). Une solution forte de l'EDS (1.1) partant à l'instant t est un processus $(X_t)_t$ progressif tel que l'on ait :

$$\int_t^s \mid b(u,X_u) \mid du + \int_t^s \mid \sigma(u,X_u) \mid^2 du < +\infty, \quad p.s., \forall t \leq s \in [0,T],$$

et que la relation

$$X_s = X_t + \int_t^s b(u,X_u)du + \int_t^s \sigma(u,X_u)dB_u \ , \tag{1.2}$$

soit vraie p.s.

Notons qu'une solution forte de l'EDS (1.1) est un processus continu.

L'existence et l'unicité d'une solution forte à l'EDS (1.1) est assurée par les conditions de Lipschitz et croissance linéaire suivantes : il existe une constante finie K tels que pour tout $t \in [0,T]$, $\omega \in \Omega$, x, $y \in \mathbb{R}$:

$$\mid b(t,x,w) - b(t,y,w) \mid + \mid \sigma(t,x,w) - \sigma(t,y,w) \mid \leq K \mid x-y \mid . \tag{1.3}$$

$$\mid b(t,x,w) \mid + \mid \sigma(t,x,w) \mid \leq K(1+ \mid x \mid). \tag{1.4}$$

Théorème 1.1.1. *Sous les conditions* (1.3), (1.4), *il existe pour tout* $t \in [0,T]$, *une solution forte à l'EDS* (3.1) *partant à l'instant* t. *De plus, pour tout* ξ $\mathcal{F}_t$ *-mesurable à valeurs dans* $\mathbb{R}$, *tel que* $E(\mid \xi \mid^2 < +\infty)$, *il ya unicité d'une solution forte* X *partant à l'instant* t *de* ξ, *i.e.*

$X_t = \xi$. L'unicité est trajectorielle (ou indistinguable). De plus cette solution est de carré intégrable et vérifie :

Pour tout $T > t$, il existe une constante C_T tel que :

$$E\left(\sup_{t \leq s \leq T} | X_s |^2\right) \leq C_T \ (1 + E(| \xi |^2)).$$

Formule de Feynman-Kac

Dans ce paragraphe, on considère l'EDS (1.1) avec des coefficients déterministes $b(t, x)$ et $\sigma(t, x)$. Pour tout $t \in \mathbb{T}$, on introduit l'opérateur (déterministe) différentiel du second-ordre :

$$(\mathcal{L}_t \varphi)(x) = b(t, x).D_x\varphi(x) + \frac{1}{2} tr \left(\sigma(t, x)\sigma'(t, x)D_x^2\varphi(x)\right), \quad \varphi \in C^2(\mathbb{R}^n).$$

$\mathcal{L}_t$ est appelé générateur infinitésimal de la diffusion (1.1). Si X est une solution de l'EDS (1.1), $v(t, x)$ une fonction (réelle) de classe $C^{1,2}$ sur $\mathbb{T} \times \mathbb{R}^n$ et $r(t, x)$ une fonction continue sur $\mathbb{T} \times \mathbb{R}^d$, on a d'après la formule d'Itô :

$$M_t = e^{-\int_0^t r(s,X_s)ds} v(t, X_t) - \int_0^t e^{-\int_0^s r(u,X_u)du} \left(\frac{\partial v}{\partial t} + \mathcal{L}_s v - rv\right)(s, X_s)ds. \quad (1.5)$$

$$= v(0, X_0) + \int_0^t e^{-\int_0^s r(u,X_u)du} D_x v(s, X_s)'\sigma(s, X_s)dW_s. \quad (1.6)$$

Le processus M est donc une martingale locale continue.

On se place sur un horizon fini $\mathbb{T} = [0, T]$ et on considère le problème d'équation aux dérivées partielles (EDP) linéaire parabolique de Cauchy :

$$rv - \frac{\partial v}{\partial t} - \mathcal{L}_t v = f, \quad sur \ \ [0, T[\times\mathbb{R}^n. \quad (1.7)$$

$$v(T, .) = g \,, \quad sur \quad \mathbb{R}^n, \quad (1.8)$$

où f (resp. g) est une fonction continue de $[0.T] \times \mathbb{R}^n$ (resp. $\mathbb{R}$) dans $\mathbb{R}$. On suppose aussi que la fonction r est positive. On donne ici une version simple du théorème de représentation de Feynman-Kac. On rappelle que $X^{t,x}$ désigne la solution de la diffusion (1.1) partant à l'instant t de x.

Représentation de Feynman-Kac

Théorème 1.1.2. *Soit v une fonction $C^{1,2}([0,T[\times\mathbb{R}^d)\cap C^0([0,T]\times\mathbb{R}^d)$ à dérivée en x bornée et solution du problème de Cauchy* (1.7) – (1.8). *Alors v admet la représentation :*

$$v(t,x)=E\left[\int_t^T e^{-\int_t^s r(u,X_u^{t,x})du} f\left(s,X_s^{t,x}\right)ds+e^{-\int_t^T r(u,X_u^{t,x})du} g(X_T^{t,x})\right], \tag{1.9}$$

pour $(t,x)\in[0,T]\times\mathbb{R}^d$.

L'application du théorème précédent requiert l'existence d'une solution régulière v au problème de Cauchy (1.7) – (1.8). Ce genre de résultats s'obtient typiquement sous une hypothèse d'uniforme ellipticité de l'opérateur $\mathcal{L}_t$:

$$\exists\varepsilon>0,\ \forall(t,x)\in[0,T]\times\mathbb{R}^n,\ \forall y\in\mathbb{R}^n,\ y'\sigma\sigma'(t,x)y\geq\varepsilon\mid y\mid^2. \tag{1.10}$$

1.2 Calcul stochastique

Définition 1.2.1. Soit $(\Omega,\mathcal{F},P)$ un espace de probabilité.

1- Une filtration est une famille $(\mathcal{F}_t, t\in\mathbb{R}^+)$ de sous tribus de $\mathcal{F}$ tel que $\mathcal{F}_s\subset\mathcal{F}_t, \forall t\geq s\geq 0$.

2- Un processus (X_t) est adapté à $(\mathcal{F}_t)$ si X_t est $\mathcal{F}_t$ -mesurable $\forall t\geq 0$.

3- Un processus $(X_t)_{t\in\mathbb{R}^+}$ est dit progressif (par rapport à $(\mathcal{F}_t)_t$) si pour tout $t\in\mathbb{R}^+$, l'application $(s,\omega)\to X_s(\omega)$ est mesurable sur $[0,t]\times\Omega$ muni de la tribu produit $B([0,t])\otimes\mathcal{F}_t$.

4- Un processus $(X_t)_{t\in T}$ est dit prévisible (par rapport à $\mathcal{F}$) si l'application $(t,\omega)\to X_s(\omega)$ est mesurable sur $T\times\ \Omega$ muni de la tribu engendrée par les processus $\mathcal{F}$-adaptés et continus.

5- Un processus $(X_t)_{t\in T}$ est dit optionnel (par rapport à $\mathcal{F}$) si l'application $(t,\omega)\to X_s(\omega)$ est mesurable sur $T\times\Omega$ muni de la tribu engendrée par les processus $\mathcal{F}$ -adaptés et càd-làg.

Soit $(\Omega,\mathcal{F},P)$ un espace de probabilité et $(\mathcal{F}_t)_{t\in\mathbb{R}^+}$ une filtration de cet espace.

Définition 1.2.2. Une famille adaptée $(M_t)_{t\in\mathbb{R}^+}$de variable aléatoire dans L^1(espace de Lebesgue : c'est l'espace des fonctions intégrables) est :

♣ Une martingale si, $\forall t\geq s,\ E[M_t/\mathcal{F}_s]=M_s$.

♣ Une surmartingale si, $\forall t\geq s,\ E[M_t/\mathcal{F}_s]\leq M_s$.

♣ Une sous-martingale si, $\forall t\geq s,\ E[M_t/\mathcal{F}_s]\geq M_s$.

Remarque 1.2.1. Une martingale (M_t) vérifie $\forall t \geq 0$ $E[M_t] = E[M_0]$.

On a alors un résultat concernant les intégrales stochastiques.
Le processus $\left(\int_0^t X_s dB_s\right)_{t\in\mathbb{R}^+}$ est une $\mathcal{F}_t-$ martingale continue, de plus on a :

$$E\left(\int_0^t X_s dB_s\right) = 0.$$

Définition 1.2.3. (Martingale locale). Soit $(X_t)_t$ un processus adapté. On dit que $(X_t)_t$ est martingale locale s'il existe une suite de temps d'arrêt $(\tau_n)_{n\geq 1}$ telle que $\lim_{n\to+\infty} \tau_n = +\infty$ p.s. et le processus arrêté X^{τ_n} est une martingale pour tout n.

Propriété de markov des solutions des EDS

La propriété de Markov pour un processus $(X_t)_{t\geq 0}$ signifie que le comportement future de ce processus après t dépend uniquement de X_t et non de ce qui s'est passé avant t.
Mathématiquement, on dira qu'un processus $(X_t)_{t\geq 0}$ vérifie la propriété de Markov par rapport à une filtration $(\mathcal{F}_t)_{t\geq 0}$ pour laquelle il est adapté, si pour tout fonction f borélienne bornée et pour tout s et t, tels que $s \leq t$:

$$E\left(f(X_t)/\mathcal{F}_s\right) = E(f(X_t)/X_s).$$

Nous allons énoncer dans ce paragraphe la propriété de markov pour une solution de (1.1).
On notera $(X_s^{t,x}, s \geq t)$ la solution de l'équation (1.1) partant de x à l'instant t et $X^x = X^{0,x}$ la solution de l'équation partant de x à l'instant $0.X^{t,x}$ vérifie pour $s \geq t$:

$$X_s^{t,x} = x + \int_t^s b\left(u, X_u^{t,x}\right) du + \int_t^s \sigma(u, X_u^{t,x}) dB_u.$$

A priori, $X^{t,x}$ est défini pour tout (t,x) presque sûrement. On peut cependant, sous les hypothèses du théorème (1.1.1), construire un processus dépendant de (t,x,s) qui est P-p.s. Continu en ces trois variables et tel que $X_s^{t,x}$ soit solution de l'équation précédente.
La propriété de Markov prend dans ce cas la forme suivante :

Théorème 1.2.1. *Soit $(X_t)_{t\geq 0}$ une solution de* (1.1). *C'est un processus de Markov par rapport à la filtration $(\mathcal{F}_t)_{t\geq 0}$ du mouvement brownien. Plus précisément, on a, pour toute fonction borélienne bornée f :*

$$P - p.s.\ E(f(X_t)/\mathcal{F}_s) = \Phi(X_s) \quad \text{où} \quad \Phi(x) = E(f(X_t^{s,x})).$$

Définition 1.2.4. (mouvement brownien standard). Un mouvement brownien standard (m.b.s) est un processus aléatoire à temps continu $(B_t, t \in \mathbb{R}_+)$ tels que :

1- $B_0 = 0$ p.s.

2- (B_t) est à accroissements indépendants et stationnaires.

3- $B_t \sim \mathcal{N}(0, t), \forall t > 0$.

4- (B_t) est à trajectoires continues.

Définition 1.2.5. (Temps d'arrêt). Soit $(\mathcal{F}_t, t \in \mathbb{R}_+)$ une filtration.

- Un temps d'arrêt T par rapport à $(\mathcal{F}_t)$ est une v.a. T à valeurs dans $\mathbb{R}_+ \cup \{+\infty\}$ telle que $\{T \leq t\} \in \mathcal{F}_t$, $\forall t \in \mathbb{R}_+$.
- On définit :

$$\mathcal{F}_t = \{A \in \mathcal{F} : A \cap \{T \leq t\} \in \mathcal{F}_t,\ \forall t \in \mathbb{R}_+\}.$$

- Si (X_t) est un processus adapté à la filtration $(\mathcal{F}_t)$, on pose $X_T(\omega) = X_{T(\omega)}(\omega)$.

Théorème 1.2.2. *(Théorème d'arrêt). Soient (M_t) une martingale continue (par rapport à $(\mathcal{F}_t)$) et T_1, T_2 deux temps d'arrêt tels que $0 \leq T_1(\omega) \leq T_2(\omega) \leq K < \infty$, $\forall \omega \in \Omega$. Alors*

$$E\left(M_{T_1}/F_{T_1}\right) = M_{T_1} \ \text{ p.s et donc } \ E\left(M_{T_2}\right) = E\left(M_{T_1}\right).$$

En particulier, si $0 \leq T(\omega) \leq K$, $\forall \omega$, alors $E(M_T) = E(M_0)$.

1.3 Formule d'Itô

La formule d'Itô est un des outils essentiels pour les applications de l'intégrale stochastique à différents problèmes.

Considérons le processus stochastique $(X_t)_t \in [0, T]$ à valeurs réelles ayant la forme :

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dB_s\ , \ (0 \leq t \leq T), \tag{1.11}$$

où $(b_t)_t \in [0,T]$ et $(\sigma_t)_t \in [0,T]$ sont deux processus stochastiques à valeurs progressivement mesurables et tels que :

$$\int_0^T |b_s|\, ds \quad < \quad +\infty.$$
$$\int_0^T |\sigma_s|^2\, ds \quad < \quad +\infty.$$

Théorème 1.3.1. *(Formule d'Itô). Soit f une fonction deux fois dérivable, et soit $(X_t)_t \in [0,T]$ ayant la forme* (1.11)*, alors on a :*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)dX_s + \frac{1}{2}\int_0^t f''(X_s)\, d\langle X, X\rangle_S\,, \tag{1.12}$$

où :

$$\langle X, X\rangle_t = \int_0^t \sigma_s^2 ds.$$

1.4 Equations différentielles partielles

Equation parabblique

Dans le cas d'un horizon fini, sous les hypothèses de régularité de b, σ, u, la fonction valeur v définie par :

$$v(t,x) = \min_{P(.)\in\mathcal{P}_{ad}} J(t,x,P(.))\,,$$

est solution de l'équation d'HJB parabolique :

$$\begin{cases} \frac{\partial v}{\partial t}(t,x) + \min\limits_{P\in\mathcal{P}_{ad}}\left(A^P v(t,x) + u(t,x,P)\right) = 0 & \text{dans } [0,T[\times\mathbb{R}^n, \\ v(T,x) = g(x). \end{cases} \tag{1.13}$$

Si l'espace d'état est un ouvert $\mathcal{O}$ de $\mathbb{R}^n$, la fonction valeur

$$v(t,x) = \min_{P(.)\in\mathcal{P}_{ad}} E\left(\int_t^{\mathrm{T}\wedge T} u(s, X_s, P_s)\, ds + f(\mathrm{T}, X_{\mathrm{T}})\, 1_{\mathrm{T}<T} + g(X_T)\, 1_{T<\mathrm{T}}/X_t = x\right), \tag{1.14}$$

où le coût f est à croissance polynomiale est solution de l'équation d'HJB

$$\left\{\begin{array}{l} \frac{\partial v}{\partial t}(t,x)+\min\limits_{P\in\mathcal{P}_{ad}}\left(A^{P}v(t,x)+u(t,x,P)\right)=0 \quad \text{dans} \quad [0.T[\times\mathcal{O}, \\ \\ v(T,x)=g(x), \end{array}\right. \tag{1.15}$$

avec la condition au limite de Dirichet

$$v\left(t,x\right)=f\left(t,x\right) \quad \text{sur} \quad [0,T[\times\partial\mathcal{O}.$$

Si le processus est réfléchi sur la frontière, la fonction valeur

$$v\left(t,x\right)=\min_{P(.)\in\mathcal{P}_{ad}}E\left(\int_{t}^{T}u\left(s,X_{s},P_{s}\right)ds+\int_{t}^{T}f\left(s,X_{s}\right)d\xi_{s}+g\left(X_{T}\right)/X_{t}=x\right), \tag{1.16}$$

est solution de l'équation d'HJB (1.15) avec la condition au limite de Neumann

$$\frac{\partial v\left(t,x\right)}{\partial n}=f\left(t,x\right) \quad \text{sur} \quad [0,T[\times\partial\mathcal{O}. \tag{1.17}$$

Equation elliptique

On se place dans le cas d'un horizon infini et on suppose que les fonctions b, σ, et u satisfont les hypothéses de régularité. On suppose de plus qu'il existe $c>0$ telle que :

$$\sum_{i=1}^{n}a_{ij}\left(x,P\right)\eta_{i}\eta_{j}\geq c\left|\eta\right|^{2}, \ \forall x\in\mathcal{O}, \ \eta\in\mathbb{R}^{n}, \ P\in\mathcal{P}_{ad}. \tag{1.18}$$

Alors, la fonction valeur v définie par :

$$v\left(x\right)=\min_{P(.)\in\mathcal{P}_{ad}}J\left(X,p\left(.\right)\right),$$

est solution de l'équation d'HJB elliptique :

$$-\lambda v\left(x\right)+\min_{P\in\mathcal{P}_{ad}}\left(A^{P}v\left(x\right)+u\left(x,P\right)\right)=0 \qquad \text{dans} \ \mathbb{R}^{n}. \tag{1.19}$$

Si $\mathcal{P}_{ad}$ est compact, le minimum dans (1.19) est toujours atteint. Sinon des conditions supplémentaires sont nécessaires.

Si l'espace d'état est un ouvert $\mathcal{O}$ de $\mathbb{R}^{n}$ et que le processus est arrêté sur la frontière, la fonction valeur est définie par :

$$v(x) = \min_{P(.)\in\mathcal{P}_{ad}} J(x,P),$$

avec

$$J(x,P) = E\left(\int_0^{\tau} e^{-\lambda t} u(X_t,P_t)\,dt + e^{-\lambda\tau} f(X_\tau)/X_0 = x\right), \tag{1.20}$$

où

$$\tau(\omega) = \inf\{t, X_t(\omega) \notin \mathcal{O}\},$$

est le premier instant où le processus X_t sort du domaine $\mathcal{O}$ et f est le coût d'arrêt. La fonction $v(x)$ satisfait l'équation d'HJB

$$-\lambda v(x) + \min_{P\in\mathcal{P}_{ad}} \left(A^P v(x) + u(x,P)\right) = 0 \qquad \text{dans } \mathcal{O}, \tag{1.21}$$

avec la condition au limite de Dirichlet

$$v(x) = f(x) \quad \text{sur } \partial\mathcal{O}. \tag{1.22}$$

Si le processus est réfléchi sur la frontière, l'équation d'évolution du processus s'écrit

$$dX_t = b(X_t,P_t)\,dt + \sigma(X_t,P_t)\,dW_t - n_{X_t} d\xi_t\,, \tag{1.23}$$

ξ_t est un processus strictement croissant lorsque $X_t \in \partial\mathcal{O}$ et $d\xi_t = 0$ si $X_t \in \mathcal{O}$, n_x est la normale extérieure à la frontière $\partial\mathcal{O}$ de $\mathcal{O}$ en x.
La fonction valeur, définie par :

$$v(x) = \min_{P(.)\in\mathcal{P}_{ad}} E\left(\int_0^{+\infty} e^{-\lambda t} u(X_t,P_t)\,dt + \int_0^{+\infty} e^{-\lambda t} f(X_t)\,d\xi_t/X_0 = x\right),$$

où f est le coût associé à la réflexion du processus, satisfait l'équation d'HJB (1.21) avec la condition au limite de Neumann

$$\frac{\partial v}{\partial n} = f \quad \text{sur } \partial\mathcal{O}. \tag{1.24}$$

La condition (1.18) est la condition d'uniforme ellipticité. Elle implique l'existence d'une solution forte de l'équation d'HJB, c'est à dire une fonction de classe $\mathcal{D}^2$ qui vérifie l'équation en tout point. L'unicité de la solution est assurée dans le cas d'un ouvert $\mathcal{O}$ borné. Il n'ya pas unicité quand l'équation est définie dans un ouvert non borné. Parmi toutes les solutions, la fonction valeur se trouve alors être celle qui ne croit pas trop rapidement lorsque $|x| \to \infty$. Lorsque (1.18) n'est pas vérifiée, l'équation est dite dégénérée et dans ce cas, il n'ya pas de solution forte à l'équation d'HJB.

La résolution de l'équation d'HJB permet de définir en tout point x un contrôle optimal $P(x)$ qui ne dépend que de x : c'est un contrôle en "feedback", c'est à dire que c'est le contrôle que l'on doit utiliser lorsqu'on se trouve au point x. Il ne dépend donc que de l'état du systéme et non pas de la variable t.

1.5 Mesure et intégration

Proposition 1.5.1. *(Théorème de la convergence dominée de Lebesgue). Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{F}, P)$ à valeurs réelles intégrables telle qu'il existe une variable aléatoire X et que $\lim\limits_{n\to+\infty} X_n = X$ P-p.s. S'il existe une variable aléatoire Y telle que $|X_n| \leq Y$ P-p.s. pour tout $n \in \mathbb{N}\backslash\{0\}$ et que $E(Y) < \infty$, alors X est intégrable et on a :*

$$\lim_{n\to\infty} E(X_n) = E(X).$$

Proposition 1.5.2. *(Inégalité de Jensen). Soit X une variable aléatoire intégrable définie sur un espace probabilisé $(\Omega, \mathcal{F}, P)$ à valeurs réeles. Soit g une fonction convexe définie sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$. Alors on a :*

$$g(E(X)) \leq E(g(X)).$$

1.6 Analyse convexe

Définition 1.6.1. (ensembe convexe). On dit que l'ensemble $\mathcal{C} \subset \mathbb{H}$ ($\mathbb{H}$ désigne un espace de Hilbert réel) est convexe si :

$$\forall (x,y) \in \mathcal{C} \times \mathcal{C},\ \forall t \in [0,1], \quad tx + (1-t)y \in \mathcal{C},$$

autrement dit, $\mathcal{C}$ est convexe s'il contient tout "segment" reliant deux quelconques de ses points.

Définition 1.6.2. (Fonction convexe(resp. concave)). On dit la fonction $J : \mathcal{C} \subset \mathbb{H} \to \mathbb{R} \cup \{+\infty\}$ est convexe (resp. concave) si $\mathcal{C}$ est convexe (resp. concave) et si :

$$\forall (x,y) \in \mathcal{C} \times \mathcal{C},\ \forall t \in [0,1], \quad J(tx + (1-t)y) \leq (resp. \text{ concave } \geq) tJ(x) + (1-t)J(y).$$

Définition 1.6.3. (Fonction strictement convexe). On dit que la fonction $J : \mathcal{C} \to \mathbb{R} \cup \{+\infty\}$ est strictement convexe si $\mathcal{C}$ est convexe et si :

$$\forall (x,y) \in \mathcal{C} \times \mathcal{C} \quad avec \ \ x \neq y,\ \forall t \in]0,1[\qquad J(tx + (1-t)y) < tJ(x) + (1-t)J(y).$$

Chapitre 2

Contrôle de processus de diffusion et principe de la programmation dynamique

La théorie du contrôle stochastique a de nombreuses applications en gestion et en finance. En effet, dans ces domaines, on considère des systèmes dynamiques (c'est à dire évoluant au cours du temps) en avenir incertain et sur lesquels on peut agir en vue d'optimiser un certain critère économique.

Pour décrire un problème de contrôle stochastique, il est important de préciser quelle est l'information disponible à tout instant. Plusieurs situations sont possibles :

- Le "contrôleur" n'a aucune information pendant l'opération du système. Dans ce cas, il choisit un contrôle qui est fonction du temps. Ces contrôles sont appelés en "boucle ouverte".

-Le contrôleur connait l'état du système à chaque instant. C'est le cas de l'observation (ou information complète).

-Le contrôleur a une connaissance partielle de l'état du système. C'est le cas de l'observation incomplète.

2.1 Contrôle de processus de diffusion

On considère un modèle de contrôle où l'état du système est gouverné par l'équation différentielle stochastique (EDS) à valeurs dans $\mathbb{R}^n$

$$dX_s = b\left(X_s, \alpha_s\right) ds + \sigma\left(X_s, \alpha_s\right) dW_s. \tag{2.1}$$

où W est un mouvement brownien d-dimentionnel sur un espace de probabilité filtré $(\Omega, \mathcal{F}, F = (\mathcal{F}_t)_{t>0}, P)$ satisfaisant les conditions habituelles. Plus généralement, on peut aussi considérer des coefficients $b(t,x,a)$ et $\sigma(t,x,a)$ dépendants du temps t. Mais dans le cas des problèmes à horizon infini décrits ci-desous, il est important de supposer que ces coefficients ne dépendent pas du temps afin d'avoir la stationnarité du problème et une fonction valeur indépendante du temps.

Le contrôle $\alpha = (\alpha_s)$ est un processus progressif (par rapport à F) et à valeurs dans A, sous espace de $\mathbb{R}^m$.

Les fonctions mesurables $b : \mathbb{R}^n \times A \to \mathbb{R}^n$ et $\sigma : \mathbb{R}^n \times A \to \mathbb{R}^{n\times d}$ satisfont une condition de Lipschitz uniforme en A : $\exists K \geq 0$, $\forall\ x,\ y \in \mathbb{R}^n$, $\forall a \in A$.

$$|b(x,a) - b(y,a)| + |\sigma(x,a) - \sigma(y,a)| \leq K\,|x-y|\,. \tag{2.2}$$

Dans la suite pour $0 \leq t \leq T \leq +\infty$, on note $\mathcal{T}_{t,T}$ l'ensemble des temps d'arrêts à valeurs dans $[t,T]$. Lorsque $t=0$ et $T=+\infty$. On note simplement $\mathcal{T}$=$\mathcal{T}_{o,+\infty}$.

2.1.1 Problème à horizon fini

On fixe un horizon fini $0 < T < +\infty$. On note par $\mathcal{A}$ l'ensemble des processus de contrôle α tel que :

$$E\left[\int_0^T |b(0,\alpha_t)|^2 + |\sigma(0,\alpha_t)|^2\, dt\right] < +\infty. \tag{2.3}$$

Le point $x=0$ est une valeur arbitraire de la diffusion et si ce point n'est pas dans le support de la diffusion, on peut choisir n'importe quelle autre valeur dans ce support. Les conditions (2.2) et (2.3) assurent pour tout $\alpha \in \mathcal{A}$ et pour toute condition initiale $(t,x) \in [0,T] \times \mathbb{R}^n$, l'existence et l'unicité d'une solution forte à l'EDS (2.1) partant de x en $s=t$. On note alors par $\{X_s^{t,x}, t \leq s \leq T\}$ cette solution qui est p.s à trajectoires continues. On rappelle aussi que sous ces conditions sur b, σ et α, on a :

$$E\left[\sup_{t\leq s\leq T} \left|X_s^{t,x}\right|^2\right] < +\infty. \tag{2.4}$$

$$\lim_{h\to 0^+} E\left[\sup_{s\in[t,t+h]} \left|X_s^{t,x} - x\right|^2\right] = 0. \tag{2.5}$$

• Critère de minimisation

Soient $f : [0,T] \times \mathbb{R}^n \times A \to \mathbb{R}$ et $g : \mathbb{R}^n \to \mathbb{R}$ des fonctions mesurables. On suppose que :

(Hg) (i) g est borné inférieurment

où (ii) g est à croissance quadratique : $|g(x)| \leq C\left(1+|x|^2\right), \forall x \in \mathbb{R}^n$, pour une constante C indépendante de x. Pour $(t,x) \in [0,T] \times \mathbb{R}^n$, on note par $\mathcal{A}(t,x)$ le sous-ensemble des contrôles α de $\mathcal{A}$ tel que :

$$E\left[\int_t^T \left|f\left(s, X^{t,x}, \alpha_s\right)\right| ds\right] < +\infty.$$

On peut alors définir sous **(Hg)** la fonction de coût :

$$J(t,x,\alpha) = E\left[\int_t^T f(s, X_s^{t,x}, \alpha_s) ds + g\left(X_T^{t,x}\right)\right].$$

Pour tous $(t,x) \in [0.T] \times \mathbb{R}^n$ et $\alpha \in \mathcal{A}(t,x)$. L'objectif étant de minimiser cette fonction coût, on introduit la fonction valeur :

$$v(t,x) = \inf_{\alpha \in \mathcal{A}(t,x)} J\left(t,x,\alpha\right). \qquad (2.7)$$

-Pour un état initial $(t,x) \in [0,T] \times \mathbb{R}^n$, on dit que $\hat{\alpha} \in \mathcal{A}(t,x)$ est un contrôle optimal si $v(t,x) = J(t,x,\hat{\alpha})$.

-un processus de contrôle α de la forme $\alpha_s = a(s, X_s^{t,x})$ pour une fonction mesurable a de $[0,T] \times \mathbb{R}^n$ dans A, est appelé contrôle markovien.

Dans la suite, on supposera implicitement que la fonction valeur v est mesurable en ses arguments.

Remarque 2.1.1. Lorsque f est à croissance quadratique en x , i.e. il existe une constante positive C et une fonction positive $\kappa : A \to \mathbb{R}_+$ telles que :

$$|f(t,x,a)| \leq C\left(1+|x|^2\right) + \kappa(a), \ \forall (t,x,a) \in [0,T] \times \mathbb{R}^n \times A, \qquad (2.8)$$

alors l'estimation (2.4) montre que pour tout $(t,x) \in [0,T] \times \mathbb{R}^n$, pour tout contrôle constant $\alpha = a$ dans A :

$$E\left[\int_t^T \left|f(s, X_s^{t,x}, a)\right| ds\right] < +\infty,$$

ainsi, les contrôles constantes dans A sont dans $\mathcal{A}(t,x)$. De plus, si il existe une constante positive C telle que $\kappa(a) \leq C\left(1 + |b(0,a)|^2 + |\sigma(0,a)|^2\right)$, pour tout a dans A, alors les conditions (2.3) et (2.4) montrent que pour tout $(t,x) \in [0,T] \times \mathbb{R}^n$, pour tout contrôle $\alpha \in \mathcal{A}$:

$$E\left[\int_t^T \left|f(s, X_s^{t,x}, \alpha_s)\right| ds\right] < +\infty.$$

Autrement dit, dans ce cas, $\mathcal{A}(t,x) = \mathcal{A}$.

2.1.2 Problème à horizon infini

On note par $\mathcal{A}_o$ l'ensemble des processus de contrôle α tel que :

$$E\left[\int_0^T |b(0,\alpha_t)|^2 + |\sigma(0,\alpha_t)|^2 dt\right] < +\infty, \quad \forall T > 0. \qquad (2.9)$$

Etant donnés une condition initiale $t = 0$ et $x \in \mathbb{R}^n$ et un contrôle $\alpha \in \mathcal{A}_0$, il existe alors une unique solution forte, noté $\{X_s^x, s \geq o\}$, de (2.1) partant de x en $t = 0$. On rappelle aussi qu'on a l'estimation suivante :

$$E\left[|X_s^x|^2\right] \leq C|x|^2 + Ce^{C_s} E\left[\int_0^s |x|^2 + |b(0,\alpha_u)|^2 + |\sigma(0,\alpha_u)|^2 du\right], \qquad (2.10)$$

pour une constante C indépondante de s, x et α.

- Critère de minimisation

Soit $\beta > 0$ et $f : \mathbb{R}^n \times A \to \mathbb{R}$ une fonction mesurable. Pour $x \in \mathbb{R}^n$, on note par $\mathcal{A}(x)$ le sous-ensemble des contrôles α de $\mathcal{A}_0$ tel que :

$$E\left[\int_0^{+\infty} e^{-\beta_s} |f(X_s^x, \alpha_s)| ds\right] < +\infty. \qquad (2.11)$$

On définit alors la fonction coût :

$$J(x,\alpha) = E\left[\int_0^{+\infty} e^{-\beta_s} f(X_s^x, \alpha_s) ds\right],$$

pour $x \in \mathbb{R}^n$ et $\alpha \in \mathcal{A}(x)$, et la fonction valeur :

$$v(x) = \inf_{\alpha \in \mathcal{A}(x)} J(x, \alpha). \qquad (2.12)$$

-pour un état initial $x \in \mathbb{R}^{n,}$ on dit que $\hat{\alpha} \in \mathcal{A}(x)$ est un contrôle optimal si $v(t, x) = J(t, x, \hat{\alpha})$.
-Un processus de contrôle α de la forme $\alpha_s = a(s, X_s^{t,x})$ pour une fonction mesurable a de $\mathbb{R}_+ \times \mathbb{R}^n$ dans A est appelé contrôle markovien.
Il est important de supposer ici que la fonction $f(x, a)$ ne dépend pas du temps pour avoir la stationnarité du problème, i.e. la fonction valeur ne dépend pas de la date initiale à laquelle on considère le problème d'optimisation.

Remarque 2.1.2. Lorsque f est à croissance quadratique en x, i.e. il existe une constante positive C et une fonction positive $\kappa : A \to \mathbb{R}_+$ telles que :

$$|f(x, a)| \leq C\left(1 + |x|^2\right) + \kappa(a), \quad \forall (x, a) \in \mathbb{R}^n \times A \qquad (2.13)$$

alors l'estimation (2.10) montre que pour $\beta > 0$ assez grand, pour tout $x \in \mathbb{R}^n$, $a \in A$:

$$E\left[\int_o^{+\infty} e^{-\beta_s} \left|f(X_s^x, a)\right| ds\right] < +\infty.$$

Autrement dit, les contrôles constants dans A appartiennent à $\mathcal{A}(x)$.

2.2 Principe de la programmation dynamique

Le principe de la programmation dynamique (PPD) est un principe fondamental pour la théorie du contrôle stochastique. Dans le contexte de contrôle de processus de diffusion décrit au paragraphe précédent, et même plus généralement pour des contrôles de processus de Markov, il s'énonce ainsi :

Théorème ***(principe de la programmation dynamique)***

On va étudiier les deux cas suivantes :

1) $Horizon\ fini$

Soit $(t, x) \in [0, T] \times \mathbb{R}^n$. Alors on a :

$$v(t,x) = \inf_{\alpha \in \mathcal{A}(t,x)} \inf_{\theta \in \mathcal{T}_{t,T}} E\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + v(\theta, X_\theta^{t,x})\right]. \quad (2.14)$$

$$= \inf_{\alpha \in \mathcal{A}(t,x)} \sup_{\theta \in \mathcal{T}_{t,T}} E\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + v(\theta, X_\theta^{t,x})\right]. \quad (2.15)$$

2) $Horizon\ infini$

Soit $x \in \mathbb{R}^n$. Alors on a :

$$v(t,x) = \inf_{\alpha \in \mathcal{A}(x)} \inf_{\theta \in \mathcal{T}} E\left[\int_0^\theta e^{-\beta_s} f(X_s^x, \alpha_s)ds + e^{-\beta\theta} v(X_\theta^x)\right], \quad (2.16)$$

$$\inf_{\alpha \in \mathcal{A}(x)} \sup_{\theta \in \mathcal{T}} E\left[\int_0^\theta e^{-\beta_s} f(X_s^x, \alpha_s)ds + e^{-\beta\theta} v(X_\theta^x)\right], \quad (2.17)$$

avec la convonction que $e^{-\beta\theta(\omega)} = 0$ lorsque $\theta(\omega) = +\infty$.

Remarque 2.2.1. Le principe de la programmation dynamique énoncé ci-dessus peut se formuler de manière équivalente, dans le cas à horizon fini.

(i) *pour tout* $\alpha \in \mathcal{A}(t,x)$ et $\theta \in \mathcal{T}_{t,T}$:

$$v(t,x) \leq E\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + v(\theta, X_\theta^{t,x})\right]. \quad (2.18)$$

(ii) *pour tout* $\delta > 0$, il existe $\alpha \in \mathcal{A}(t,x)$ tel que pour tout $\theta \in \mathcal{T}_{t,T}$:

$$v(t,x) + \delta \geq E\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + v(\theta, X_\theta^{t,x})\right]. \quad (2.19)$$

C'est une version plus forte que la version traditionnelle du principe de la programmation dynamique, qui s'écrit en horizon fini

$$v(t,x) = \inf_{\alpha \in \mathcal{A}(t,x)} E\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + v(\theta, X_\theta^{t,x})\right]. \quad (2.20)$$

Pour tout temps d'arrêt $\theta \in \mathcal{T}_{t,T}$. On a une remarque analogue dans le cas à horizon infini. L'idée intuitive de ce principe est qu'un contrôle optimal $\hat{\alpha}$ sur $[t,T]$ peut être recollé en deux contrôles optimaux, l'un sur $[t,\theta]$ et l'autre sur $[\theta,T]$, et ceci quel que soit le temps d'arrêt θ. La preuve rigoureuse de ce résultat dans ce contexte stochastique est très technique. Nous donnons ici une preuve formelle de ce principe.

Preuve formelle du PPD

On concidère le cas de problème à horizon fini.

1. Etant donné un contrôle $\alpha \in \mathcal{A}(t,x)$, on a par unicité du flot de l'EDS gouvernant X, la structure markovienne :

$$X_s^{t,x} = X_s^{\theta, X_\theta^{t,x}}, \ s \geq \theta,$$

où θ est un temps d'arrêt à valeurs dans $[t,T]$. Par la loi des espérances conditionnelles itérées, on a alors :

$$J(t,x,\alpha) = E\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + J(\theta, X_\theta^{t,x}, \alpha)\right],$$

d'où puisque $J(.,.,\alpha) \geq \nu$ et comme θ est quelconque dans $\mathcal{T}_{t,T}$

$$\begin{aligned} J(t,x,\alpha) &\geq \sup_{\theta\in\mathcal{T}_{t,T}} E\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + v(\theta, X_\theta^{t,x})\right] \\ &\geq \inf_{\alpha\in\mathcal{A}(t,x)} \sup_{\theta\in\mathcal{T}_{t,T}} E\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + v(\theta, X_\theta^{t,x})\right]. \end{aligned}$$

En passant à l'infimum sur α dans le terme de gauche, on obtient l'inégalité :

$$v(t,x) \geq \inf_{\alpha\in\mathcal{A}(t,x)} \sup_{\theta\in\mathcal{T}_{t,T}} E\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + v(\theta, X_\theta^{t,x})\right]. \tag{2.21}$$

2. On considère pour tout $\varepsilon > 0$ et $\theta \in \mathcal{T}_{t,T}$, un contrôle ε-optimal α^ε de $v(\theta, X_\theta^{t,x})$

$$J(\theta, X_\theta^{t,x}) \leq v(\theta, X_\theta^{t,x}) + \varepsilon.$$

Etant donné $\alpha \in \mathcal{A}(t,x)$, on définit le processus :

$$\hat{\alpha}_s = \begin{cases} \alpha_s, s \in [0,\theta] \\ \alpha_s^{\varepsilon}, s \in [\theta,T] \end{cases}.$$

Le point délicat est de vérifier que $\hat{\alpha}$ est progressivement mesurable et alors bien dans $\mathcal{A}(t,x)$. Dans ce cas, on a :

$$\begin{aligned} v(t,x) &\leq J(t,x,\hat{\alpha}) = E\left[\int_t^\theta f(s,X_s^{t,x},\alpha_s)ds + J(\theta,X_\theta^{t,x},\alpha^{\varepsilon})\right] \\ &\leq E\left[\int_t^\theta f(s,X_s^{t,x},\alpha_s)ds + v(\theta,X_\theta^{t,x})\right] + \varepsilon. \end{aligned}$$

Ceci étant valable pour tout $\alpha \in \mathcal{A}(t,x)$, $\theta \in \mathcal{T}_{t,T}$ et $\varepsilon > 0$, on a l'inégalité :

$$v(t,x) \leq \inf_{\alpha \in \mathcal{A}(t,x)} \inf_{\theta \in \mathcal{T}_{t,T}} E\left[\int_t^\theta f(s,X_s^{t,x},\alpha_s)ds + v(\theta,X_\theta^{t,x})\right]. \tag{2.23}$$

En combinant les deux inégalités (2.21) et (2.22), on obtient le résultat voulu.

Chapitre 3

Equation d'Hamilton-Jacobi-Bellman

L'équation d'Hamilton -Jacobi -Bellman (HJB) est la version infinitésimale du principe de la programmation dynamique : elle décrit le comportement local de la fonction valeur $v(t,x)$ lorsqu'on fait tendre le temps d'arrêt θ dans (2.20) vers t.
Dans cette section, nous dérivons formellement l'équation d'HJB en suposant que la fonction valeur v est suffisamment régulière.

3.1 Dérivation formelle de HJB

Problème à horizon fini

Considérons le temps $\theta = t+h$ et un contrôle constant $\alpha_s = a$, avec a arbitraire dans A, dans la relation (2.18) de la programmation dynamique :

$$v(t,x) \leq \; E\left[\int_t^{t+h} f(s, X_s^{t,x}, a)ds + v(t+h, X_{t+h}^{t,x})\right]. \qquad (3.1)$$

En supposant que v est suffisamment régulière, on a par la formule d'Itô entre t et $t+h$:

$$v\left(t+h, X_{t+h}^{t,x}\right) = v(t,x) + \int_t^{t+h} \left(\frac{\partial v}{\partial t} + \mathcal{L}^a v\right)\left(s, X_s^{t,x}\right) ds + martingale\ (locale),$$

où $\mathcal{L}^a$ est l'opérateur associé à la diffusion (2.1) pour le contrôle constant a et défini par :

$$\mathcal{L}^a v = b(x,a).D_x v + \frac{1}{2} tr\left(\sigma(x,a)\sigma'(x,a)D_x^2 v\right).$$

En substituant dans (2.23), on obtient alors :

$$0 \leq E\left[\int_t^{t+h} (\frac{\partial v}{\partial t} + \mathcal{L}^a v)(s, X_s^{t,x}) + f(s, X_s^{t,x}, a)ds\right].$$

En divisant par h et en faisant tendre h vers 0, on a :

$$0 \leq \frac{\partial v}{\partial t}(t,x) + \mathcal{L}^a v(t,x) + f(t,x,a).$$

Ceci étant valable pour tout $a \in A$, on a l'inégalité :

$$-\frac{\partial v}{\partial t}(t,x) + \sup_{\alpha \in A}\left[-\mathcal{L}^a v(t,x) - f(t,x,a)\right] \leq 0. \tag{3.2}$$

D'autre part, supposons que α^*est un contrôle optimal. Alors dans (2.20), on a :

$$v(t,x) = E\left[\int_t^{t+h} f(s, X_s^*, \alpha_s^*)ds + v(t+h, X_{t+h}^*)\right],$$

où X^*est l'état du système solution de (2.1) partant de x en t avec le contrôle α^*. Par un argument similaire et avec des conditions de régularités sur v, on obtient :

$$-\frac{\partial v}{\partial t}(t,x) - \mathcal{L}^{\alpha_t^*} v(t,x) - f(t,x,\alpha_t^*) = 0, \tag{3.3}$$

ce qui combiné avec (2.1) suggère que v doit satisfaire :

$$-\frac{\partial v}{\partial t}(t,x) + \sup_{\alpha \in A}[-\mathcal{L}^a v(t,x) - f(t,x,a)] = 0, \ \forall (t,x) \in [0,T[\times \mathbb{R}^n, \tag{3.4}$$

si le supremum ci-dessus en a est fini. Nous verrons plus tard comment traiter le cas où ce supremum est infini, ce qui peut intervenir lorsque l'espace des contrôles A est non bornée. On reécrit souvent cette équation aux dérivées partielles (EDP) sous la forme :

$$-\frac{\partial v}{\partial t}(t,x) + H(t,x,D_x v(t,x), D_x^2 v(t,x)) = 0, \quad \forall (t,x) \in [0,T] \times \mathbb{R}^n, \tag{3.5}$$

où pour $(t,x,p,M) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times S_n$ (S_n est l'ensmble des matrices $n \times n$ symétrique)

$$H(t,x,p,M) = \sup_{a \in A}\left[-b(x,a).p - \frac{1}{2}tr(\sigma\sigma'(x,a)M) - f(t,x,a)\right].$$

Cette fonction H est appelée Hamiltonien du problème de contrôle considéré.

Cette équation (3.4) est appelée équation de la programmation dynamique ou équation de Hamilton-Jacobi-Bellman (HJB). A cette équation aux dérivées partielles, il faut ajouter la condition terminale :

$$v(T,x) = g(x), \forall x \in \mathbb{R}^n, \qquad (3.6)$$

qui résulte immédiatement de la définition (2.7) de la fonction valeur v.

Remarque 3.1.1. 1) Lorsque l'ensemble des contrôles est réduit à un singleton $\{a_0\}$, c'est à dire qu'il n'y a pas de contrôle sur l'état du système, l'EDP d' HJB se réduit au problème d'EDP linéaire de Cauchy :

$$-\frac{\partial v}{\partial t}(t,x) - \mathcal{L}^{a_0} v(t,x) = f(t,x,a_0), \ \forall (t,x) \in [0,T[\times \mathbb{R}^n \ . \quad (3.7)$$

$$v(T,x) = g(x), \ \forall x \in \mathbb{R}^n \ . \qquad (3.8)$$

2) L'argument d'optimalité de la programmation dynamique suggère que si l'on peut trouver un contrôle $\alpha^*(t,x)$ telle que :

$$\sup_{\alpha \in A} [-\mathcal{L}^a v(t,x) - f(x,a)] = -\mathcal{L}^{\alpha^*(t,x)} v(t,x) - f(x,\alpha^*(t,x)),$$

c'est à dire que :

$$\alpha^*(t,x) \in \arg \ \min_{a \in A} [\mathcal{L}^a v(t,x) + f(x,a)] \, ,$$

alors on aura :

$$-\frac{\partial v}{\partial t} - \mathcal{L}^{\alpha^*(t,x)} v(t,x) - f(x,\alpha^*(t,x)) = 0,$$

et donc

$$v(t,x) = E\left[\int_t^T f(X_s^*, \alpha^*(s,X_s^*))ds + g(X_T^*) \right],$$

où X^* est solution de l'équation différentielle stochastique :

$$\begin{aligned} dX_s^* &= b(X_s^*, \alpha^*(s,X_s^*)) + \sigma(X_s^*, \alpha^*(s,X_s^*))dW_s, \ t \le s \le T, \\ X_t^* &= x, \end{aligned}$$

et α^*est un contrôle optimal Markovien.

Problème à horizon infini

En reprenant les mêmes arguments que dans le cas d'un problème à horizon infini, on dérive l'équation d'HJB pour la fonction valeur définie en (2.12) :

$$\beta v(x) + \sup_{a \in A} [-\mathcal{L}^{\alpha} v(x) - f(x,a)] = 0, \ \forall x \in \mathbb{R}^n,$$

qu'on reécrit aussi sous la forme :

$$\beta v + H(x, Dv(x), D^2 v(x)) = 0, \ \forall x \in \mathbb{R}^n,$$

où pour $(x, p, M) \in \mathbb{R}^n \times \mathbb{R}^n \times S_n$:

$$H(x,p,M) = \sup_{a \in A} \left[-b(x,a).p - \frac{1}{2} tr(\sigma\sigma'(x,a)M) - f(x,a) \right].$$

Remarques-Extensions

1. Les résultats dans le cas horizon fini s'étendent aisément lorsque la fonction de coût J à minimiser a la forme plus générale suivante :

$$J(t,x,\alpha) = E\left[\int_t^T \Gamma(t,s) f(s, X_s^{t,x}, \alpha_s) ds + \Gamma(T) g(X_T^{t,x}) \right],$$

où

$$\Gamma(t,s) = \exp\left(-\int_t^s \beta(u, X_u^{t,x}, \alpha_u) du \right), \ t \le s \le T,$$

et $\beta(.,.,a)$ est une fonction continue positive sur $[0,T[\times\mathbb{R}^n$ pour tout $a \in A$.

Dans ce cas d'Hamiltonien associé au problème de contrôle stochastique est :

$$H(t,v,x,p,M) = \sup_{a \in A} \left[\beta(t,x,a)v - b(x,a).p - \frac{1}{2} tr(\sigma(x,a)\sigma'(x,a)M) - f(t,x,a) \right].$$

2. Lorsque l'espace des contrôles A est non borné, l'Hamiltonien

$$H(t,x,p,M) = \sup_{a\in A}\left[-b(x,a).p - \frac{1}{2}tr(\sigma\sigma'(x,a)M) - f(t,x,a)\right].$$

peut prendre la valeur $+\infty$ dans un certain domaine de (t,x,p,M). Plus précisément, supposons qu'il existe une fonction continue $G(t,x,p,M)$ sur $[0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times S_n$ telle que :

$$H(t,x,p,M) < +\infty \Leftrightarrow G(t,x.p,M) \leq 0.$$

Alors d'après le raisonnement conduisant à l'équation d'HJB (3.4), on doit avoir :

$$G(t,x,D_xv(t,x),D_x^2v(t,x)) \leq 0, \qquad (3.9)$$

et

$$-\frac{\partial v}{\partial t}(t,x) + H(t,x,D_xv(t,x),D_x^2v(t,x)) \leq 0\ . \qquad (3.10)$$

De plus si l'inégalité (3.8) est stricte en un point $(t,x)\in[0,T]\times\mathbb{R}^n$, alors il existe un voisinage de $(t,x,D_xv(t,x),D_x^2v(t,x))$ en lequel H est fini. Alors, formellement, le contrôle devrait être attient au voisinage de (t,x) et par un raisonnement analogue à (3.4) on doit avoir l'égalité dans (3.9). On obtient ainsi une inéquation variationnelle d'HJB pour la programmation dynamique

$$\max\left[-\frac{\partial v}{\partial t}(t,x) + H(t,x,D_xv(t,x),D_x^2v(t,x)), G(t,x,D_xv(t,x),D_x^2v(t,x))\right] = 0.$$

On dit que le problème de contrôle est singulier. Un cas typique est losque le contrôle intervient de manière linéaire dans la dynamique du système et dans la fonction de coût. Par exemple, dans le cas unidimensionnel $n=1$, $A=\mathbb{R}_-$, et

$$b(x,a) = a + \hat{b}(x),\ \ \sigma(x,a) = \hat{\sigma}(x),$$

$$f(t,x,a) = a + \hat{f}\ (t,x),$$

alors

$$H(t,x,p,M) = \begin{cases} -\hat{b}(x)p - \frac{1}{2}\hat{\sigma}(x)^2 M - \hat{f}\ (t,x) & si \quad p+1 \leq 0 \\ +\infty & si \quad p+1 > 0 \end{cases}.$$

L'inéquation variationnelle d'HJB s'écrit ainsi :

$$\max\left[-\frac{\partial v}{\partial t}(t,x) - \hat{b}(x)\frac{\partial v}{\partial x}(t,x) - \frac{1}{2}\hat{\sigma}(x)^2\frac{\partial^2 v}{\partial x^2}(t,x), \frac{\partial v}{\partial x}(t,x)+1\right] = 0.$$

3. Lorsqu'on étudie un problème de maximisation

$$v(t,x) = \sup_{\alpha \in \mathcal{A}(t,x)} E\left[\int_t^T f(s, X_s^{t,x}, \alpha_s)ds + g(X_T^{t,x})\right],$$

on peut se ramener à un problème de minimisation en considérant la fonction valeur $-v$. Ceci revient alors à considérer l'Hamiltonien

$$H(t,x,p,M) = \inf_{a \in A(t,x)}\left[-b(x,a).p - \frac{1}{2}tr(\sigma(x,a)\sigma'(x,a)M) - f(t,x,a)\right],$$

avec une équation d'HJB :

$$-\frac{\partial v}{\partial t}(t,x) + H(t,x,D_x v(t,x), D_x^2 v(t,x)) = 0.$$

Lorsque H peut prendre la valeur $-\infty$ en supposant qu'il existe une fonction continue $G(t,x,p,M)$ sur $[0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times S_n$ telle que :

$$H(t,x,p,M) > -\infty \Leftrightarrow G(t,x.p,M) \geq 0,$$

l'inéquation variationnelle d' HJB s'écrit :

$$\min[-\frac{\partial v}{\partial t}(t,x) + H(t,x,D_x v(t,x), D_x^2 v(t,x)), G(t,x,D_x v(t,x), D_x^2 v(t,x))] = 0.$$

On a une remarque analogue dans le cas d'un problème de maximisation en horizon infini.

3.2 Théorème de vérification

L'etape la plus important dans la programmation dynamique consiste à montrer, étant donnée une solution régulière à l'équation d'HJB, que ce condidat, sous des conditions suffissantes, coincide avec la fonction valeur. Ce résultat est appelé théorème de vérification et permet aussi d'obtenir un contrôle optimal. Il repose essentiellement sur la formule d'Itô.

Théorème 3.2.1. (*horizon fini*)

Soit $\omega \in C^{1,2}([0,T[\times\mathbb{R}^n)\cap C^0([0,T]\times\mathbb{R}^n)$ à croissance quadratique, i.e. il existe une constante C telle que :

$$|\omega(t,x)| \leq C\left(1+|x|^2\right), \quad \forall (t,x)\in[0,T]\times\mathbb{R}^n.$$

(i) Supposons que :

$$\begin{aligned} -\frac{\partial\omega}{\partial t}(t,x)+\sup_{a\in A}\left[-\mathcal{L}^{\alpha}\omega(t,x)-f(t,x,a)\right] &\leq 0, \quad (t,x)\in[0,T[\times\mathbb{R}^n && (3.11)\\ \omega(T,x) &\leq g(x), \ x\in\mathbb{R}^n. && (3.12)\end{aligned}$$

Alors $\omega\leq v$ sur $[0,T]\times\mathbb{R}^n$.

(ii) De plus supposons que $\omega(T.) = g$, et pour tout $(t,x)\in[0,T[\times\mathbb{R}^n$, il existe $\hat{\alpha}(t,x)$ mesurable à valeurs dans A tel que :

$$-\frac{\partial\omega}{\partial t}(t,x)+\sup_{a\in A}\left[-\mathcal{L}^{\alpha}\omega(t,x)-f(t,x,a)\right] = -\frac{\partial\omega}{\partial t}(t,x)-\mathcal{L}^{\hat{\alpha}(t,x)}\omega(t,x)-f(t,x,\hat{\alpha}(t,x))=0,$$

l'EDS :

$$dX_s = b(X_s,\hat{\alpha}(s,X_s))ds+\sigma(X_s,\hat{\alpha}(s,X_s))dW_s,$$

admette une solution, notée $\hat{X}_s^{t,x}$, étant donnée une condition initiale $X_t = x$, et $\left\{\hat{\alpha}(s,\hat{X}_s^{t,x})\ t\leq s\leq T\right\}\in\mathcal{A}(t,x)$. Alors :

$$\omega = v \quad \text{sur} \quad [0,T]\times\mathbb{R}^n,$$

et $\hat{\alpha}$ est un contrôle optimal Markovien.

Preuve. (i) Puisque $\omega \in C^{1,2}([0,T[\times \mathbb{R}^n)$, on a pour tout $(t,x) \in [0,T[\times\mathbb{R}^n$, $\alpha \in \mathcal{A}(t,x)$, $s \in [t,T[$, et pour tout temps d'arrêt τ à valeurs dans $[t,+\infty[$, par la formule d'Itô :

$$\omega(s\wedge\tau, X_{s\wedge\tau}^{t,x}) = \omega(t,x)+\int_t^{s\wedge\tau} \frac{\partial\omega}{\partial t}(u,X_u^{t,x})+L^{\alpha_u}\omega(u,X_u^{t,x})du+\int_t^{s\wedge\tau} D_x\omega(u,X_u^{t,x})'\sigma(X_u^{t,x},\alpha_u)dW_u.$$

On choisit $\tau = \tau_n = inf\left\{s \geq t : \int_t^s \left|D_x\omega(u,X_u^{t,x})'\sigma(X_u^{t,x},\alpha_u)\right|^2 du \geq n\right\}$ en notant que $\tau_n \to +\infty$ quand n tend vers l'infini.
Le processus arrêté $\left\{\int_t^{s\wedge\tau_n} D_x\omega(u,X_u^{t,x})'\sigma(X_u^{t,x},\alpha_u)dW_u, t \leq s \leq T\right\}$ est donc une martingale et on a en prenant l'espérance :

$$E\left[\omega(s\wedge\tau_n, X_{s\wedge\tau_n}^{t,x})\right] = \omega(t,x) + E\left[\int_t^{s\wedge\tau_n} \frac{\partial\omega}{\partial t}(u,X_u^{t,x}) + L^{\alpha_u}\omega(u,X_u^{t,x})du\right].$$

Puisque ω satisfait (3.10), on a :

$$\frac{\partial\omega}{\partial t}(u,X_u^{t,x}) + L^{\alpha_u}\omega(u,X_u^{t,x}) + f(X_u^{t,x},\alpha_u) \geq 0, \ \forall\alpha \in \mathcal{A}(t,x),$$

d'où :

$$E\left[\omega(s\wedge\tau_n, X_{s\wedge\tau_n}^{t,x})\right] \geq \omega(t,x) - E\left[\int_t^{s\wedge\tau_n} f(X_u^{t,x},\alpha_u)du\right], \ \forall\alpha \in \mathcal{A}(t,x). \tag{3.13}$$

On a :

$$\left|\int_t^{s\wedge\tau_n} f(X_u^{t,x},\alpha_u)du\right| \leq \int_t^T \left|f(X_u^{t,x},\alpha_u)\right| du,$$

et le terme de droite est intégrable d'après la condition d'intégrabilité sur $\mathcal{A}(t,x)$. Comme ω est à croissance quadratique, on a :

$$\left|\omega(s\wedge\tau_n, X_{s\wedge\tau_n}^{t,x})\right| \leq C\left(1+\sup_{s\in[t,T]} | X_s^{t,x} |^2\right),$$

et le terme de droite est intégrable d'après (2.4). On peut donc appliquer le théorème de convergence dominée et faire tendre n vers l'infini dans (3.12) :

$$E[\omega(s,X_s^{t,x})] \geq \omega(t,x) - E[\int_t^s f(X_u^{t,x},\alpha_u)du], \ \forall\alpha \in A(t,x).$$

Comme ω est continue sur $[0,T] \times \mathbb{R}^n$, en faisant tendre s vers T, on a par le théoème de convergence dominée et en utilisant aussi (3.11) :

$$E\left[g(X_T^{t,x})\right] \geq \omega(t,x) - E\left[\int_t^T f(X_u^{t,x}, \alpha_u)du\right], \ \forall \alpha \in \mathcal{A}(t,x),$$

et donc

(ii) On applique la formule d'Itô à $\omega(u, \hat{X}{}_u^{t,x})$ entre $t \in [0,T[$ et $s \in [t,T[$:

$$E\left[\omega(s,\hat{X}_s^{t,x})\right] = \omega(t,x) + E\left[\int_t^s \frac{\partial \omega}{\partial t}(u,\hat{X}_u^{t,x}) + L^{\hat{\alpha}(u,\hat{X}_u^{t,x})}\omega(u,\hat{X}_u^{t,x})du\right].$$

Or par définition de $\hat{\alpha}(t,x)$, on a :

$$-\frac{\partial \omega}{\partial t} - L^{\hat{\alpha}(t,x)}\omega(t,x) - f(t,x,\hat{\alpha}(t,x)) = 0,$$

d'où :

$$E\left[\omega(s,\hat{X}_s^{t,x})\right] = \omega(t,x) - E\left[\int_t^s f(\hat{X}_u^{t,x}, \hat{\alpha}(u,\hat{X}_u^{t,x}))du\right].$$

En faisant tendre s vers T, on obtient ainsi :

$$\begin{aligned} \omega(t,x) &= E\left[\int_t^T f(\hat{X}_u^{t,x}, \hat{\alpha}(u,\hat{X}_u^{t,x}))du + g(\hat{X}_T^{t,x})\right], \\ &= J(t,x,\hat{\alpha}). \end{aligned}$$

Donc

$$\omega(t,x) = J(t,x,\hat{\alpha}) \geq \nu(t,x).$$

Et finalement $\omega = \nu$ avec $\hat{\alpha}$ comme contrôle optimale markovien. □

Remarque 3.2.1. Dans le cas particulier où l'espace des contrôle A est réduit à un singleton $\{a_0\}$, ce théorème de vérification est une version du théorème de représentation de Feynman-Kac il stipule que si ω est une fonction $C^{1,2}$ $([0,T[\times\mathbb{R}^n) \cap C^0([0,T[\times\mathbb{R}^n)$ à croissance quadratique solution du problème de Cauchy (3.6) − (3.7), alors ω admet la représentation :

$$\omega(t,x) = E\left[\int_t^T f(X_s^{t,x}, a_0)ds + g(X_T^{t,x})\right].$$

Théorème 3.2.2. *(Horizon infini)*

Soit $\omega \in C^2(\mathbb{R}^n)$ à croissance quadratique.

(i) Supposons que :

$$\beta\omega(x) + \sup_{a\in A}\left[-\mathcal{L}^a\omega(x) - f(x,a)\right] \leq 0, \ x \in \mathbb{R}^n \ . \tag{3.14}$$

$$\liminf_{T\to+\infty} \ e^{-\beta T}E\left[\omega(X_T^x)\right] \leq 0, \ \forall x \in \mathbb{R}^n, \ \forall \alpha \in \mathcal{A}(x). \tag{3.15}$$

Alors

$$\omega \leq v \ \ sur \ \ \mathbb{R}^n.$$

(ii) Supposons de plus que pour tout $x \in \mathbb{R}^n$, il existe $\hat{\alpha}(x)$ mesurable à valeurs dans A tel que :

$$\begin{aligned}\beta\omega(x) + \sup_{a\in A}\left[-\mathcal{L}^a\omega(x) - f(x,a)\right] &= \beta\omega(x) - \mathcal{L}^{\hat{\alpha}(x)}\omega(x) - f(x,\hat{\alpha}(x)). \\ &= 0,\end{aligned}$$

l'EDS :

$$dX_s = b\left(X_s, \hat{\alpha}(X_s)\right)ds + \sigma\left(X_s, \hat{\alpha}(X_s)\right)dW_s,$$

admette une solution, notée $\hat{X}_s^x$, étant donnée une condition initiale $X_0 = x$, avec :

$$\lim\ \sup_{T\to+\infty} \ e^{-\beta T}E\left[\omega(\hat{X}_T^x)\right] \geq 0, \tag{3.16}$$

et telle que

$$\left\{\hat{\alpha}(\hat{X}_s^x), s \geq 0\right\} \in \mathcal{A}(x).$$

Alors

$$\omega(x) = v(x), \ \forall x \in \mathbb{R}^n,$$

et $\hat{\alpha}$ est un contrôle optimal markovien.

Preuve. (i) Soit $\omega \in C^2(\mathbb{R}^n)$ et $\alpha \in \mathcal{A}(x)$. On a par la formule d'Itô à $e^{-\beta t}\omega(X_t^x)$ entre 0 et $T \wedge \tau_n$:

$$e^{-\beta(T\wedge\tau_n)}\omega\left(X_{T\wedge\tau_n}^x\right) = \omega(x) + \int_0^{T\wedge\tau_n} e^{-\beta u}\left[L^{\alpha_u}\omega(X_u^x - \beta\omega(X_u^x))\right]du + \int_0^{T\wedge\tau_n} e^{-\beta u}D\omega\left(X_u^x\right)'\sigma(X_u^x,\alpha_u)dW_u.$$

Ici, τ_n est le temps d'arrêt :

$$\tau_n = \inf\left\{t \geq 0 \ : \ \int_0^t \left|D_x\omega(u, X_u^x)'\sigma(X_u^x,\alpha_u)\right|^2 du \geq n\right\}.$$

Comme le terme d'intégrale stochastique arrêtée est une martingale, on a en prenant l'espérance :

$$\begin{aligned} E\left[e^{-\beta(T\wedge\tau_n)}\omega(X_{T\wedge\tau_n}^x)\right] &= \omega(x) + E\left[\int_0^{T\wedge\tau_n} e^{-\beta u}(-\beta\omega + L^{\alpha_u}\omega)(X_u^x)du\right]. \\ &\geq \omega(x) - E\left[\int_0^{T\wedge\tau_n} e^{-\beta u}f(X_u^x,\alpha_u)du\right], \end{aligned} \tag{3.17}$$

d'après (3.13). Par la condition de croissance quadratique de ω et la condition d'intégrabilité (3.11), on peut appliquer le théorème de convergence dominée et faire tendre n vers l'infini :

$$E\left[e^{-\beta T}\omega(X_T^x)\right] \geq \omega(x) - E\left[\int_0^T e^{-\beta u}f(X_u^x,\alpha_u)du\right], \ \forall \alpha \in \mathcal{A}(x).$$

En faisant tendre T vers l'infini, on a d'après (3.14) et le théorème de convergence dominée :

$$\omega(x) \leq E\left[\int_0^{+\infty} e^{-\beta t}f(X_t^x,\alpha_t)dt\right], \ \forall \alpha \in \mathcal{A}(x),$$

et donc

$$\omega(x) \leq \nu(x), \ \forall x \in \mathbb{R}^n.$$

(ii) En appliquant la formule d'Itô à $e^{-\beta t}\ \omega(\hat{X}_s^x)$ et en observant que le contrôle $\hat{\alpha}$ atteint l'égalité dans (3.16), on a :

$$E\left[e^{-\beta T}\omega(\hat{X}_T^x)\right] = \omega(x) - E\left[\int_0^T e^{-\beta u} f(\hat{X}_u^x, \hat{\alpha}(\hat{X}_u^x))du\right].$$

En faisant tendre T vers l'infini et d'après (3.15), on obtient ainsi :

$$\omega(x) \geq J(x,\hat{\alpha}) = E\left[\int_0^\infty e^{-\beta u} f(\hat{X}_u^x, \hat{\alpha}(\hat{X}_u^x))du\right],$$

et donc

$$\omega(x) = \nu(x) = J(x,\hat{\alpha}).$$

□

Remarque 3.2.2. Dans le cas particulier où l'espace des contrôles A est réduit à un singleton $\{a_0\}$, ce théorème de vérification est une version du théorème de représentation de Feynman-Kac en horizon infini : il stipule que si ω est une fonction $C^2(\mathbb{R}^n)$ à croissance quadratique solution du problème d'EDP elliptique

$$\beta\omega(x) - \mathcal{L}^{a_0}\omega(x) - f(x,a_0) = 0,\ \ x \in \mathbb{R}^n,$$

$$\lim_{T\to+\infty} e^{-\beta T} E\left[\omega(X_T^x)\right] = 0,\ \ x \in \mathbb{R}^n,$$

alors ω admet la représentation

$$\omega(t,x) = E\left[\int_0^{+\infty} e^{-\beta t} f(X_t^x, a_0)dt\right].$$

Le théorème précédent suggère la stratégie suivante pour résoudre le problème de contrôle stochastique. Dans le cas horizon fini, résoudre l'EDP nonlinéaire d'Hamilton-Jacobi-Bellman :

$$-\frac{\partial\omega}{\partial t} + \sup_{a\in A}\left[-\mathcal{L}^a\omega(t,x) - f(t,x,a)\right] = 0,\ \ (t,x) \in [0,T[\times\mathbb{R}^n\ , \tag{3.18}$$

avec la condition terminale $\omega(T,x) = g(x)$. Fixer $(t,x) \in [0,T[\ \times\mathbb{R}^n$ et résoudre :

$$\sup_{a\in A}[-\mathcal{L}^a\omega(t,x)-f(x,a)],$$

comme un problème de maximum en $a\in\ A$.

On note $a^*(t,x)$ la valeur de a qui réalise le maximum.

Si cette EDP nonlinéaire avec condition terminale admet une solution régulière ω, alors ω est la fonction valeur du problème de contrôle stochastique et a^*est un contrôle optimal markovien. Cette méthode se justifie donc si l'EDP d'HJB (3.17) admet une solution $C^{1,2}$ satisfaisant les conditions d'application du théorème de vérification. Des résultats d'existence de solution régulières à des équations de type HJB paraboliques (horizon fini) ou elliptiques (horizon infini) sont établis dans Fleming et Rishel [2], Gilbarg et Trudinger[3]ou encore Krylov [5]. La condition principale imposée est une condition d'uniforme ellipticité :

il existe une canstante $c>0$ telle que

$$y'\sigma(x,a)\sigma'(x,a)y\geq c\left|y\right|^2,\forall x\ ,\ y\in\mathbb{R}^n,\ \forall a\in A.$$

Soulignons aussi dans la vérification des conditions (ii) des théorèmes 3.2.1 et 3.2.2 qu'il n'est pas toujours facile et parfois problématique d'obtenir l'existence d'une solution à l'EDS associé au candidat $\hat{\alpha}$ pour être le contrôle optimal.

3.3 Applications

3.3.1 Problème de choix de portefeuille de Merton en horizon fini

On a l'exemple suivant :

Un agent investit à chaque date t une proportion α_t de sa richesse dans un actif risqué S et $1-\alpha_t$ dans un actif sans risque S^0, avec la contrainte qu'à toute date, α_t doit être à valeurs dans A ensemble fermé convexe de $\mathbb{R}$. Son processus de richesse évolue selon l'EDS :

$$\begin{aligned} dX_t &= \frac{X_t\alpha_t}{S_t}dS_t+\frac{X_t\left(1-\alpha_t\right)}{S_t^0}dS_t^0 \\ &= X_t(\alpha_t\mu+(1-\alpha_t)r)dt+X_t\alpha_t\sigma dW_t. \end{aligned}$$

Partant d'une richesse initiale $X_t = x > 0$ au temps t, l'agent veut maximiser l'espérance de l'utlité de sa richesse terminale à un horizon $T > t$. Notons par $X^{t,x}$ le processus de richesse partant de x en t. Observons que les coefficients de X ne vérifient pas stricto-sensu la condition de Lipschitz uniforme en les contrôles a. En fait, on se ramène usuellement à ce cadre en considérant le logarithme de la richesse positive. La fonction valeur du problème de maximisation est donc :

$$v(t,x) = \sup_{\alpha \in \mathcal{A}} E\left[U(X_T^{t,x})\right], \ (t,x) \in [0,T] \times \mathbb{R}_+,$$

où $\mathcal{A}$ est l'ensemble des processus α progressifs à valeurs dans A et tels que :

$$E\left[\int_0^T |\alpha_s|^2 \, ds\right] < +\infty.$$

La fonction d'utilité U est croissante et concave sur $\mathbb{R}_+$. Vérifions que $v(t,.)$ est croissante et concave en x.

Soit $0 < x \leq y$ et α un processus de contrôle dans $\mathcal{A}$.

Notons :

$$Z_s = X_s^{t,x} - X_{s.}^{t,y}.$$

Alors :

$$dZ_s = Z_s\left[(\alpha_s \mu + (1-\alpha_s)r)ds + \alpha_s \sigma dW_s\right], \ Z_t = y - x \geq 0,$$

et donc :

$$Z_s \geq 0,$$

ou encore :

$$X_s^{t,y} \geq X_s^{t,x} \text{ pour tout } s \geq t.$$

Puisque U est croissante, on a $U(X_T^{t,x}) \leq U(X_T^{t,y})$ d'où :

$$E\left[U(X_T^{t,x})\right] \leq E\left[U(X_T^{t,y})\right] \leq v(t,y), \ \forall \alpha \in \mathcal{A}_t,$$

et donc :

$$v(t,x) \leq v(t,y).$$

Soit : $0 < x_1$, x_2, α^1, α^2 deux processus de contrôle dans $\mathcal{A}$ et $\lambda \in [0,1]$.

Posons : $x_\lambda = \lambda x_1 + (1-\lambda)x_2$, X^{t,x_1} le processus de richesse partant de x_1 et controlé par α^1, X^{t,x_2} le processus de richesse partant de x_2 et controlé par α^2.

Posons :

$$\alpha_s^\lambda = \frac{\lambda X_s^{t,x_1}\alpha_s^1 + (1-\lambda)X_s^{t,x_2}\alpha_s^2}{\lambda X_s^{t,x_1} + (1-\lambda)X_s^{t,x_2}}.$$

Notons que par convexité de A, le processus $\alpha^\lambda \in \mathcal{A}$. De plus, d'après la structure linéaire de l'évolution de l'équation de la richesse, le processus :

$$X^\lambda = \lambda X^{t,x_1} + (1-\lambda)\; X^{t,x_2},$$

est gouverné par :

$$\begin{aligned} dX_s^\lambda &= X_s^\lambda \left(\alpha_s^\lambda \mu + (1-\alpha_s^\lambda) r\right) ds + X_s^\lambda \alpha_s^\lambda \sigma dW_s,\; s \geq t. \\ X_t^\lambda &= x_\lambda. \end{aligned}$$

Ceci prouve que $\lambda X^{t,x_1} + (1-\lambda)X^{t,x_2}$ est un processus de richesse partant de x_λ en t et contrôlé par α^λ.

D'après la concavité de la fonction U, on a :

$$U\left(\lambda X_T^{t,x_1} + (1-\lambda)X_T^{t,x_2}\right) \geq \lambda U\left(X_T^{t,x_1}\right) + (1-\lambda)\, U\left(X_T^{t,x_2}\right).$$

D'où :

$$v\left(\lambda x_1 + (1-\lambda)x_2\right) \geq \lambda E\left[U(X_T^{t,x_1})\right] + (1-\lambda)E\left[U(X_T^{t,x_2})\right].$$

Et ceci pour tout α^1, α^2 dans $\mathcal{A}$. On en déduit que :

$$v\left(\lambda x_1 + (1-\lambda)x_2\right) \geq \lambda \nu(x_1) + (1-\lambda)v(x_2).$$

En fait, on voit que si U est strictement concave et s'il existe un contrôle optimal, alors les arguments ci-dessus montrent que la fonction v est aussi strictement concave en x.
On va donc chercher à résoudre l'équation de Bellman :

$$-\frac{\partial \omega}{\partial t} + \inf_{a \in A} \left[-\mathcal{L}^a w(t,x) \right] = 0, \tag{3.19}$$

avec la condition terminale

$$\omega(T,x) = U(x), \ x \in \mathbb{R}_+. \tag{3.20}$$

Ici :

$$\mathcal{L}^a w(t,x) = x(a\mu + (1-a)r)\frac{\partial w}{\partial x} + \frac{1}{2}x^2 a^2 \sigma^2 \frac{\partial^2 w}{\partial x^2}.$$

Le problème $(3.19)-(3.20)$ n'a en général pas de solution explicite pour une fonction d'utilité U quelconque.
Cependant, dans le cas particulier d'une fonction puissance :

$$U(x) = x^p, \ x \geq 0, \ p < 1.$$

On peut résoudre explicitement ce problème. Cherchons une solution de la forme :

$$w(t,x) = \Phi(t) x^p.$$

En substituant dans $(3.19)-(3.20)$, on obtient que Φ satisfait :

$$\begin{aligned} -\Phi'(t) + \rho\Phi(t) &= 0. \\ \Phi(T) &= 1. \end{aligned}$$

où

$$\rho = \inf_{a \in A} \left[-ap(\mu - r) - pr + \frac{1}{2}a^2 p(1-p)\sigma^2 \right]. \tag{3.21}$$

On obtient alors :

$$\Phi(t) = \exp(-\rho(T-t)).$$

Ainsi, la fonction donnée par :

$$w(t,x) = \exp(-\rho(T-t))x^p, \ (t,x) \in [0,T] \times \mathbb{R}_+, \tag{3.22}$$

est régulière, strictement croissante et strictement concave, et est solution de (3.19) – (3.20). De plus, la fonction :

$$a \in A \to -ap\,(\mu - r) - rp + \frac{1}{2}a^2 p\,(1-p)\,\sigma^2,$$

est strictement convexe sur l'ensemble convexe A donc atteint son minimum en $\hat{\alpha}$ constant. Par construction, $\hat{\alpha}$ atteint l'infimum de $\inf_{a \in A} [-\mathcal{L}^a w(t,x)]$.

De plus, l'équation de la richesse associée au contrôle constant $\hat{\alpha}$:

$$dX_t = X_t(\hat{\alpha}\mu + (1-\hat{\alpha})r)dt + X_t\hat{\alpha}\sigma dW_t,$$

admet bien une unique solution, étant donnée une condition initiale. Ceci prouve donc finalement d'aprés le théorème de vérification, que la fonction valeur du problème de maximisation est donnée par (3.22) et que la proportion optimale de richesse à investir dans l'actif risqué est constante et donnée par $\hat{\alpha}$.

Notons que lorsque $A = \mathbb{R}$, on a :

$$\hat{a} = \frac{\mu - r}{\sigma^2(1-p)}, \tag{3.23}$$

et

$$\rho = -\frac{(\mu - r)^2}{2\sigma^2}\frac{p}{1-p} - rp.$$

3.3.2 Un modèle de production et consommation en horizon infini

On considère le modèle suivant d'une unité de production. La valeur K_t de son capital à la date t varie selon le taux d'investissement I_t en capital et le prix S_t par unité de capital :

$$dK_t = K_t \frac{dS_t}{S_t} + I_t dt.$$

La dette L_t de cette unité de production évolue selon le taux d'intérêt r, la consommation C_t et le taux de productivité P_t du capital :

$$dL_t = rL_t dt - \frac{K_t}{S_t} dP_t + (I_t + C_t)dt.$$

On choisit un modèle d'évolution de $(Y_t = \ln S_t, P_t)$ selon :

$$\begin{aligned} dY_t &= \left(\mu - \frac{\sigma_1^2}{2}\right) dt + \sigma_1 dW_t^1, \\ dP_t &= bdt + \sigma_2 dW_t^2, \end{aligned}$$

où (W^1, W^2) est un mouvement brownien de dimension 2 sur un espace de probabilité filtré $(\Omega, \mathcal{F}, F = (\mathcal{F}_t)_{t\geq 0}, P)$ et μ, b, σ_1, σ_2 sont des constantes, $\sigma_1, \sigma_2 > 0$.

La valeur nette de l'unité de production est :

$$X_t = K_t - L_t.$$

On impose les contraintes :

$$K_t \geq 0,\ C_t \geq 0,\ X_t \geq 0,\ t \geq 0.$$

On note par :

$$k_t = K_t / X_t,$$

et

$$c_t = C_t / X_t.$$

Les variables de contrôle d'investissement et de consommation.

La dynamique du système contrôlé est donc gouvernée par :

$$dX_t = X_t\left[k_t\left(\mu - r + be^{-Y_t}\right) + (r - c_t)\right]dt + k_t X_t \sigma_1 dW_t^1 + k_t X_t e^{-Y_t}\sigma_2 dW_t^2 \tag{3.24}$$

$$dY_t = \left(\mu - \frac{\sigma_1^2}{2}\right)dt + \sigma_1 dW_t^1 \tag{3.25}$$

Etant donné un facteur d'actualisation $\beta > 0$ et une fonction d'utilité :

$$U(C) = \frac{C^\gamma}{\gamma},\ C \geq 0 \text{ où } 0 < \gamma < 1.$$

On note par $\mathcal{A}(x, y)$ l'ensemble des processus de contrôle (k, c) progressifs à valeurs dans $\mathbb{R}_+ \times \mathbb{R}_+$tels que :

$$\int_0^T k_t^2 dt + \int_0^T c_t^2 dt < +\infty,\ ps,\ \forall T > 0,$$

$$E\left[\int_0^{+\infty} e^{-\beta t} U(c_t X_t^{x,y})dt\right] < \infty,$$

où $(X^{x,y}, Y^y)$ est la solution de l'EDS $(3.24) - (3.25)$ partant de (x, y) en $t = 0$.

L'objectif est de déterminer l'investissement et la consommation optimale de l'unité de production.

On cherche à résoudre le problème de contrôle stochastique en horizon infini :

$$v(t, x) = \sup_{(k,c)\in\mathcal{A}(x,y)} E\left[\int_0^{+\infty} e^{-\beta t} U\left(c_t X_t^{x,y}\right)dt\right] \tag{3.26}$$

L'équation d'Hamilton-Jacobi -Bellman associée est :

$$\begin{aligned}
&\beta v - \left(\mu - \frac{\sigma_1^2}{2}\right)\frac{\partial v}{\partial y} - rx\frac{\partial v}{\partial x} - \frac{\sigma_1^2}{2}\frac{\partial^2 v}{\partial y^2} + \inf_{c\geq 0}\left[cx\frac{\partial v}{\partial x} - U(cx)\right] + \\
&\inf_{k\geq 0}\left[-k\left(\mu - r + be^{-y}\right)x\frac{\partial v}{\partial x} - \frac{1}{2}k^2x^2\left(\sigma_1^2 + e^{-2y}\sigma_2^2\right)\frac{\partial^2 v}{\partial x^2} - kx\sigma_1^2\frac{\partial^2 v}{\partial x \partial y}\right] \\
= \ & 0.
\end{aligned}$$

En remarquant que $X^{x,y}$ s'exprime sous la forme :

$$X^{x,y} = x \exp\left(Z\left(y\right)\right),$$

où $Z(y)$ s'écrit comme une exponentielle de processus en fonction de (k, c) et Y^y, on en déduit que :

$$v(x,y) = x^\gamma v(1,y).$$

On cherche donc une solution de HJB (3.27) sous la forme :

$$v(x,y) = \frac{x^\gamma}{\gamma}\exp(\varphi(y)),$$

où $\varphi(y)$ est une fonction de $\mathbb{R}$ dans $\mathbb{R}$.

En substituant cette forme dans (3.27), on obtient l'équation différentielle ordinaire (EDO) que doit satisfaire φ :

$$\beta-\gamma r-\frac{\sigma_1^2}{2}\left(\varphi_{yy}+\varphi_y^2\right)-\left(\mu-\frac{\sigma_1^2}{2}\right)\varphi_y+\inf_{c\geq 0}\left[c\gamma-c^\gamma e^{-\varphi}\right]+\gamma\inf_{k\geq 0}G\left(y,\varphi_y,k\right)=0, \tag{3.28}$$

où

$$G(y,p,k) = \frac{k^2}{2}(1-\gamma)(\sigma_1^2+\sigma_2^2e^{-2y}) - k(\mu - r + be^{-y} + \sigma_1^2 p).$$

On peut montrer que pour β assez grand, il existe une unique solution régulière C^2 φ bornée à l'EDO (3.28).

Comme $G(y,p,0) = 0$, on a qu'en tout point y extrémum de φ_y, i.e. $\varphi_{yy}(y) = 0$:

$$0 \leq \beta - \gamma r - \left(\mu - \frac{\sigma_1^2}{2}\right)\varphi_y(y) - \frac{\sigma_1^2}{2}\varphi_y^2(y) - (1-\gamma)e^{\frac{\gamma\varphi(y)}{\gamma-1}}.$$

Comme φ est bornée, ceci prouve que φ_y est aussi bornée sur $\mathbb{R}$.

Par construction, la fonction positive $w(x,y) = (x^\gamma/\gamma)\ e^{\varphi(y)}$ est solution de HJB (3.27).

D'après la première partie du théorème de vérification, on en déduit que $w \geq v$.

D'autre part, considérons les fonctions :

$$\hat{k}(y) = \left(\frac{be^{-y} + \mu - r + \sigma_1^2 \varphi_y}{(1-\gamma)(\sigma_1^2 + \sigma_2^2 e^{-2y})}\right) \in \arg\min_{k \geq 0} G(y, \varphi_y, k).$$

et

$$\hat{c}(y) = \exp\left(\frac{\varphi(y)}{\gamma - 1}\right) \in \arg\min_{c \geq 0} [c\gamma - c^{\gamma} e^{-\varphi}].$$

Comme φ et φ_y sont bornées, ceci implique que les fonctions $\hat{c}$, $\hat{k}$, $e^{-y}\hat{k}$ sont aussi bornées en y.

On en déduit qu'il existe une constante $M > 0$ telle que :

$$E\left[\left|\hat{X}_t^{x,y}\right|^2\right] \leq x^2 \exp(Mt), \ \forall t > 0,$$

où on a noté par $\hat{X}^{x,y}$ la solution de (2.46) contrôlée par $\left(\hat{k}(Y_t^y), \hat{c}(Y_t^y)\right)_{t \geq 0}$.

Ceci montre que le contrôle $(\hat{k}(Y_t^y), \hat{c}(Y_t^y))_{t \geq 0}$ est dans $\mathcal{A}(x, y)$:

$$E\left[\int_0^{+\infty} e^{-\beta t} U(\hat{c}(Y_t^y)\hat{X}_t^{x,y}) dt\right] < +\infty. \tag{3.29}$$

De plus, comme φ est bornée, la fonction $\hat{c}(y)$ est bornée inférieurement par une constante strictement positive.

On en déduit l'existence d'une constante $B > 0$ telle que :

$$0 \leq e^{-\beta T} w\left(\hat{X}_T^{x,y}, Y_T^y\right) \leq Be^{-\beta T} U\left(\hat{c}\left(Y_T^y\right) \hat{X}_T^{x,y}\right).$$

Ce qui combiné avec (3.29), montre que :

$$\lim_{T \to +\infty} E\left[e^{-\beta T} w(\hat{X}_T^{x,y}, Y_T^y)\right] = 0.$$

On conclut avec la deuxième partie (ii) du théorème de vérification que $w = v$ et que $\left(\hat{k}(Y_t^y), \hat{c}(Y_t^y)\right)_{t \geq 0}$ est un contrôle optimal markovien.

3.3.3 Exemple de problème de contrôle stochastique singulier

On considère le processus contrôlé réel gouverné par :

$$dX_s = \alpha_s dW_s,$$

où α est un processus progressif à valeurs dans $A = \mathbb{R}$ et telle que :

$$E\left[\int_0^T |\alpha_s|^2 \, ds\right] < +\infty.$$

On note $\mathcal{A}$ l'ensemble de ces processus de contrôles.

Soit g une fonction mesurable positive ou à croissance linéaire sur $\mathbb{R}$, et considérons le problème de contrôle stochastique :

$$v(t,x) = \inf_{\alpha \in \mathcal{A}} E\left[g(X_T^{t,x})\right], \ (t,x) \in [0,T] \times \mathbb{R}. \tag{3.30}$$

Nous allons voir que pour un large choix de fonction g, la fonction valeur v n'est pas régulière. D'après le principe de la programmation dynamique, on a pour tout temps d'arrêt $\theta \in \mathcal{T}_{t,T}$, et tout contrôle constant $\alpha_t = a \in \mathbb{R}$:

$$v(t,x) \leq E\left[v(\theta, X_\theta^{t,x})\right] \tag{3.31}$$

Supposons que v est régulière $C^{1,2}$ et appliquons la formule d'Itô à $v\left(s, X_s^{t,x}\right)$ entre :

$$s = t \in [0,T[,$$

$$et$$

$$s = \theta = (t+h) \wedge \mathsf{T} \ où \ \mathsf{T} = \inf\left\{s \geq t : \left|X_s^{t,x} - x\right| \geq 1\right\}.$$

Alors l'intégrale stochastique apparaissant dans la formule d'Itô est une martingale arrêtée et en substituant dans (3.31), on obtient :

$$0 \leq E\left[\frac{1}{h}\int_t^{(t+h)\wedge T}\left(\frac{\partial v}{\partial t}+a^2\frac{\partial^2 v}{\partial x^2}\right)(s, X_s^{t,x})ds\right]. \tag{3.32}$$

Notons que par continuité p.s. de la trajectoire de $X_s^{t,x}$, on a que pour $h \leq \bar{h}(w)$ suffisamment petit, $\theta = t + h$ p.s.

On en déduit par le théorème de la moyenne que la variable aléatoire sous le signe espérance dans (3.32) converge p.s. vers :

$$\left(\frac{\partial v}{\partial t}+a^2\frac{\partial^2 v}{\partial x^2}\right)(t,x),$$

quand h tend vers zéro, de plus, cette variable aléatoire étant bornée par une constante indépendante de h, on peut appliquer le théorème de la convergence dominée et obtenir quand h tend vers zéro :

$$0 \leq \frac{\partial v}{\partial t}(t,x) + a^2\frac{\partial^2 v}{\partial x^2}(t,x), \ \forall (t,x) \in [0,T[\times\mathbb{R}. \tag{3.33}$$

Cette inégalité étant valable pour tout a dans $\mathbb{R}$, on doit avoir en particulier que $\frac{\partial^2 v}{\partial x^2} \geq 0$ sur $[0,T[\times\mathbb{R}$, autrement dit :

$$v(t,.) \text{ est convexe sur } \mathbb{R} \text{ pour tout } t \in [0,T[. \tag{3.34}$$

D'autre part, comme le contrôle constant nul est dans A, il est immédiat que :

$$v(t,x) \leq g(x), \ \forall (t,x) \in [0,T] \times \mathbb{R}.$$

En notant par g^{conv} l'enveloppe convexe de g, i.e. la plus grande fonction convexe minorant g, on en déduit d'après (3.34) :

$$v(t,x) \leq g^{conv}(x), \ \forall (t,x) \in [0,T[\times\mathbb{R}. \tag{3.35}$$

En utilisant $g^{conv} \leq g$, l'inégalité de Jensen et la propriété de martingale de $X_s^{t,x}$ pour $\alpha \in \mathcal{A}$, on a :

$$\begin{aligned} v(t,x) &\geq \inf_{\alpha \in \mathcal{A}_t} E\left[g^{conv}(X_T^{t,x})\right] \\ &\geq \inf_{\alpha \in \mathcal{A}_t} g^{conv}(E[X_T^{t,x}]) = g^{conv}(x). \end{aligned}$$

En combinant avec (3.35), on en déduit que :

$$v(t,x) = g^{conv}(x), \ \forall (t,x) \in [0,T[\times \mathbb{R}. \tag{3.36}$$

On aboutit à une contradiction dès lorsque la fonction g^{conv} n'est pas $C^2(\mathbb{R})$, par exemple si :

$$g(x) = \max(x-k,0) = g^{conv}(x).$$

3.4 Références bibliografique

[1] Barles. G : Solutions de viscosité des équations de Hamilton-Jacobi.1994

[2] Fleming. W et Rechel. R : deterministic and stochastic optimal control. Springer Verlag. 1975

[3] Gelbarg. D et Trudinger N : Elliptic differential eqations of second order. Springer Verlag. 1985

[4] Karatzas. I et Shreve. S : Brownian motion and stochastic calculus.Springer Verlag. 1988

[5] Krylov. N : nonlenear elliptic and parabolic equations of second order. Boston, D. Reidel.

[6] Oksendal. B : Stochastic Differential Equations.Springer Verlag.

[7] Pham. H : Optimisation et contrôle stochastique appliqués à la finance.2007

[8] Pham. H : Contrôle optimal stochastique et applications en finance,Notes de cour.

[9] Protter. P : Stochastic integration and differential equations.Springer Verlag. 1990

[10] Revuz. D : Mesures et intégration. Hermann. 1994

[11] Robert. F : Les systèmes dynamiques discrets.1995

[12] Rockafellar. R : Convex analysis. Princeton University Press. 1970